Emerging Approaches for Environmental Health Data Integration: Data on Clinical Outcomes

Atul Butte, MD, PhD
Priscilla Chan and Mark Zuckerberg Distinguished Professor
Director, Institute for Computational Health Sciences
University of California, San Francisco

atul.butte@ucsf.edu
@atulbutte
Conflicts of Interest

- **Scientific founder and advisory board membership**
  - Genstruct
  - NuMedii
  - Personalis
  - Carmenta

- **Honoraria for talks**
  - Lilly
  - Pfizer
  - Siemens
  - Bristol Myers Squibb
  - AstraZeneca
  - Roche
  - Genentech
  - Warburg Pincus

- **Past or present consultancy**
  - Lilly
  - Johnson and Johnson
  - Roche
  - NuMedii
  - Genstruct
  - Tercica
  - Ecoeos
  - Helix
  - Ansh Labs
  - Prevenda
  - Samsung
  - Assay Depot
  - Regeneron
  - Verinata
  - Pathway Diagnostics
  - Geisinger Health
  - Covance
  - Wilson Sonsini Goodrich & Rosati
  - Orrick
  - 10X Genomics
  - Medgenics
  - GNS Healthcare
  - Gerson Lehman Group
  - Coatue Management
  - Allergan
  - Astellas
  - Thomson Reuters
  - Intel
  - SAP
  - SV Angel
  - Progenity
  - Illumina

- **Speakers’ bureau**
  - None

- **Companies started by students**
  - Carmenta
  - Serendipity
  - Stimulomics
  - NunaHealth
  - Praedicat
  - MyTime
  - Flipora
  - Tumbl.in

- **Corporate Relationships**
  - Northrop Grumman
  - Aptalis
  - Northrop Grumman
  - Aptalis
The Precision Medicine Initiative

The time is right because of:
- Sequencing of the human genome
- Improved technologies for biomedical analysis
- New tools for using large datasets

The future of health begins with you.

joinallofus.org
We are building a research program of 1,000,000+ people

The mission of the All of Us Research Program is to accelerate health research and medical breakthroughs, enabling individualized prevention, treatment, and care for all of us.
NIH funds additional medical centers to expand national precision medicine research program

The awardees are:

- California Precision Medicine Consortium: University of California San Diego, with partners Cedars-Sinai Medical Center, Los Angeles; San Diego Blood Bank; University of California, Davis; University of California Health; University of California, Irvine; University of California, San Francisco; and University of Southern California, Los Angeles
Beta Testing Begins for NIH’s All of Us Research Program

June 5, 2017

I'm happy to announce that we've begun enrolling our first participants as beta testers of the All of Us Research Program. This is a major milestone in our progress to date, and one of many more to come for our startup research consortium that came together less than a year ago.

The mission of the All of Us Research Program is to accelerate health research and medical breakthroughs, enabling individualized prevention, treatment, and care for all of us.

There have been several important steps leading up to this point. Back in July of 2016, about the time I arrived at NIH, we assembled an amazing nationwide team of universities, medical centers, and technology companies to enroll participants and collect data and blood and urine samples. We've completed preliminary pilot studies and focus groups to learn from members of the public about their interests and questions concerning research participation. We've
Approved Protocol

https://allofus.nih.gov/sites/default/files/allofus-initialprotocol-v1_0.pdf

- Adults 18 and older in United States or territory
- Outreach: TV, web, print advertisements, social media
- Enrollment: Direct Volunteers or Health Care Provider Organizations (HPO)
- Electronic consenting
- Physical exam
- Participant provided information: includes geolocations
- Biospecimens: 10 ml urine, 44.5 ml blood
- EHR data from HPO recruits
- Array of sensors and wearables used later in the study
All of Us Wants Your Ideas!

Are you interested in helping researchers understand more about why people get sick or stay healthy? If so, we need your ideas to help make the All of Us Research Program the best resource it can be to support future research.

The All of Us Research Program, led by the National Institutes of Health, aims to build one of the largest, most diverse datasets of its kind to help speed up research on many different health conditions. We hope that 1 million or more people across the country will join. Participants will share information about their health, habits, and what it’s like where they live. By looking for patterns in the data, researchers may learn more about the factors that affect our health. The program will last for many years and allow us to study health over time.

See the current research protocol for more details.
AOU RESEARCH PRIORITIES USE CASES

How can we better identify and assess environmental and genetic/physiological risk factors that lead to child/teen stress?

There is tremendous pressure kids are exposed to. This results in an unusual number of children and teen age kids suffering from stress and depression, leading in some cases to devastating consequences. A study that could address to what level there are genetic predispositions to stress and depression and how it intersects with modern pressures and that will involve school, children and parents as cohorts, this could ...more »

Submitted by Lis Caler (@calar) on 29th Jan | Add your comment
<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provider Name</td>
<td>Johnson, Paul</td>
<td></td>
</tr>
<tr>
<td>Patient Name</td>
<td>Johnson, Paul</td>
<td></td>
</tr>
<tr>
<td>Address</td>
<td>5 First Street, CA</td>
<td></td>
</tr>
<tr>
<td>City</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zip</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date of Service</td>
<td>12/18/04</td>
<td></td>
</tr>
<tr>
<td>Diagnosis Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedure Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of Service</td>
<td></td>
<td></td>
</tr>
<tr>
<td>billed amount</td>
<td>7500</td>
<td></td>
</tr>
<tr>
<td>Payment Received</td>
<td>6500</td>
<td></td>
</tr>
<tr>
<td>Amount Due</td>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>Date Claim was Submitted</td>
<td>12/19/04</td>
<td></td>
</tr>
</tbody>
</table>
Sutter’s $1 Billion Boondoggle-New Electronic Records System Goes Dark

California Nurses Association Press Release, 8/27/13

Yet Another Risk to Patient Safety

A controversial electronic health records system on which Sutter corporation has said it is spending $1 billion went completely dark Monday at Sutter hospitals in Northern California exposing patients to additional risk beyond problems reported with the system in July, registered nurses reported yesterday.

Partners’ $1.2b patient data system seen as key to future

Aims for one file per person, fewer errors

How Kaiser bet $4 billion on electronic health records -- and won

Kaiser Permanente CIO Philip Fasano explains how electronic records have paid off and the health care giant’s embrace of mobile technology

InfoWorld | May 2, 2013
(2)(i) The following identifiers of the individual or of relatives, employers, or household members of the individual, are removed:

<table>
<thead>
<tr>
<th>(A) Names</th>
</tr>
</thead>
</table>
| (B) All geographic subdivisions smaller than a state, including street address, city, county, precinct, ZIP code, and their equivalent geocodes, except for the initial three digits of the ZIP code if, according to the current publicly available data from the Bureau of the Census:
   (1) The geographic unit formed by combining all ZIP codes with the same three initial digits contains more than 20,000 people; and
   (2) The initial three digits of a ZIP code for all such geographic units containing 20,000 or fewer people is changed to 000. |
| (C) All elements of dates (except year) for dates that are directly related to an individual, including birth date, admission date, discharge date, death date, and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older. |
| (D) Telephone numbers |
| (E) Fax numbers |
| (F) Email addresses |
| (G) Social security numbers |
| (H) Medical record numbers |
| (I) Health plan beneficiary numbers |
| (J) Account numbers |
| (L) Vehicle identifiers and serial numbers, including license plate numbers |
| (M) Device identifiers and serial numbers |
| (N) Web Universal Resource Locators (URLs) |
| (O) Internet Protocol (IP) addresses |
| (P) Biometric identifiers, including finger and voice prints |
| (Q) Full-face photographs and any comparable images |
| (R) Any other unique identifying number |
The next big data: clinical data

SEARCH 15 MILLION+ PATIENT RECORDS FROM THE UNIVERSITY OF CALIFORNIA WITH THE UC ReX DATA EXPLORER

ABOUT THE TOOL  ABOUT THE DATA  GET STARTED
Combining healthcare data from across the six University of California medical schools and systems
The University of California and UnitedHealth Group are teaming up to form a new accountable care organization (ACO) and clinically integrated network. As part of the 10-year strategic relationship, UC Health’s five academic medical centers will expand use of Optum’s clinically integrated network services and advanced data analytics services.
UC Health Active Patients (2 visits in last 3 years)
University of California Cancer Consortium Takes on California’s $14 Billion Killer

By Elizabeth Fernandez on September 11, 2017

University of California President Janet Napolitano announced the formation of a UC Cancer Consortium to include five of the nation’s leading academic cancer centers during a press conference in Genentech Hall at Mission Bay. Photo by Susan Merrelli
UC Health: Cancer Center Patients (2016)
Mean Age = 48
UC Clinical Data Warehouse Team

Executive Team
• Atul Butte
• Joe Bengfort
• Michael Pfeffer
• Tom Andriola
• Chris Longhurst

Steering Committee
• Irfan Chaudhry
• Mohammed Mahbouba
• Lisa Dahm
• David Dobbs
• Kent Andersen
• Ralph James
• Jennifer Holland
• Eugene Lee

ETL Team
• Albert Dugan

• Tony Choe
• Michael Sweeney
• Timothy Satterwhite
• Ayan Patel
• Niranjan Wagle
• Ralph James
• Joseph Dalton

Epic
• Kevin Ames
• Ben Jenkins
• Steve Gesualdo

Business Analyst
• Ankeeta Shukla

Data Harmonization
• Dana Ludwig
• Daniella Meeker

Data Quality
• Momeena Ali
• Jodie Nygaard

Hardware
• Sandeep Chandra
• Jeff Love
• Scott Bailey
• Kwong Law
• Pallav Saxena

Support
• Jack Stobo
• Michael Blum
• Sam Hawgood
Support

- University of California, San Francisco
- NIH: NIAID, NLM, NIGMS, NCI, NHLBI, OD; NIDDK, NHGRI, NIA, NCATS
- March of Dimes
- Juvenile Diabetes Research Foundation
- Hewlett Packard
- Howard Hughes Medical Institute
- California Institute for Regenerative Medicine
- Luke Evnin and Deann Wright (Scleroderma Research Foundation)
- Clayville Research Fund
- PhRMA Foundation
- Stanford Cancer Center, Bio-X, SPARK

- Tarangini Deshpande
- Kimayani Butte
- Sam Hawgood
- Keith Yamamoto
- Isaac Kohane

Admin and Tech Staff
- Mary Lyall
- Mounira Kenaani
- Kevin Kaier
- Boris Oskotsky
- Mae Moredo
- Ada Chen